Article

Geometry and Control of Thermodynamic Systems described by Generalized Exponential Families

Marco Favretti
Dipartimento di Matematica "Tullio Levi-Civita"
via Trieste 63, 35123 Padova, Italy; favretti@math.unipd.it

Abstract: In this paper we investigate the geometric structure and control of exponential families depending on additional parameters, called external parameters. These generalized exponential families emerges naturally when one applies the maximum entropy formalism to derive the equilibrium statistical mechanics framework. We study the associated statistical model, compute the Fisher metric and introduce a natural fibration of the parameter space over the external parameter space. The Fisher Riemannian metric allows to endow this fibration with an Ehresmann connection and to study the geometry and control of these statistical models. As an example, we show that horizontal lift of paths in the external parameter space corresponds to an isentropic evolution of the system. We apply the theory to the example of a ideal gas in a rotating rigid container. Most of the results are expressed in local coordinates; in the appendices we hint at possible global extensions of the theory.

Keywords: statistical models; Fisher metric; Ehresmann connection, exponential families

1. Introduction

The pioneering works of R.A. Fisher [1] and C. R. Rao [2] have opened the way to the application of differential geometry tools to statistics. These seminal ideas have been developed by many others scientists creating a nowaday active research field called Information Geometry or Geometric information science (see [3] and [4] for a non exhaustive list of references). The main geometric object of interest for the present paper is the nowaday classical notion of (parametric, regular) statistical model $\mathcal{S}$, a set collecting all the probability densities $p = p_z$ depending smoothly on a finite number of parameters z. The set $\mathcal{S}$ can be given the structure of a Riemannian manifold with respect to the celebrated Fisher information metric $g = g(z)$ defined as

$$g_{ij}(z) = \int_X \frac{\partial \ln p_z}{\partial z_i}\frac{\partial \ln p_z}{\partial z_j} p_z dx. \quad (1)$$

Particularly important examples are statistical models described by exponential families, i.e. of the form (sum from 1 to a over repeated indices is understood)

$$p_\theta(x) = e^{-\theta_\alpha h_\alpha(x) - \psi(\theta)}, \qquad \text{where} \quad \psi(\theta) = \ln \int_X e^{-\theta_\alpha h_\alpha(x)} dx.$$

In this paper we consider a generalization of exponential families by supposing that the random variables $h_\alpha = h_\alpha(x, u)$ depends on so-called external parameters u, which are to be distinguished from the natural parameters θ. These generalized exponential families arise naturally when one applies the Maximum Entropy formalism to equilibrium Statistical Mechanics. It is well known that when the information consist of the average values of some random variables h_α describing observables of interest for the system, the maximum entropy probability densities are exponential families. Indeed, if we introduce the Shannon entropy function for a probability density p defined on X

$$H(p) = -\int_X p \ln p dx, \qquad \int_X p dx = 1. \quad (2)$$

then the probability density that maximize H on the set of probability densities that satisfy the constraints $\langle h \rangle = \int_X h p dx = c \in \mathbb{R}^a$ has the form of an exponential family (see [5], [7]). This fact points out an important difference between the natural parameters θ and the external ones u because the formers are the Lagrange multipliers associated to the constraints when one solves the constrained extremization problem for H using Lagrange multipliers method, while the latters are parameters in the problem formulation that can be fixed or varied by an agent external to the system under consideration. Typical examples of external parameters are the magnetic or electric field applied to the system or the length of a polymer chain ([8,13]). For equilibrium statistical mechanics, a prominent object is the physical entropy of the system which for $h_\alpha = h_\alpha(x, u)$ depends on the parameters $z = (\theta, u)$ and it is defined as

$$S(\theta, u) = -\int_X p_\theta \ln p_\theta dx = \psi(\theta, u) + \theta_\alpha \langle h_\alpha \rangle.$$

The aim of this paper is twofold: on the one hand (see Section 2) we want to compute the Fisher information metric (1) in case of parametric statistical model defined by generalized exponential families and derive in this generalized setting the analog of the first principle of thermodynamics in the form of an equality between one forms (see [5], [7])

$$dS = \theta_i \Big(d\langle h_i \rangle - \langle \partial_k h_i \rangle du_k \Big).$$

On the other hand, (see Section 4) the presence in the statistical model of parameters of different kind θ and u leads in a natural way to consider the trivial fibration in the parameter space $\mathcal{Z} \times U$

$$\pi : \mathcal{Z} \times U \longrightarrow U, \qquad \pi(\theta, u) = u \tag{3}$$

Fibers of π represent standard exponential families where the external parameters have fixed values. The presence of the Fisher Riemannian metric allows to define the subspace g-orthogonal to the vertical subspace (tangent to the fibers). This geometrical framework allows to interpret the above introduced one form dS using the connection one form and to discuss globally defined geometrical notions as the horizontal lift of a path in base space, the parallel transport with respect to the connection or the holonomy of a path. It turns out that the horizontal lift of a path in the external parameter space U represent an isentropic ($dS = 0$) evolution of the system. To make the paper self contained, in Section 3 we give a brief account of the theory of Ehresmann connections on fibered bundles introducing the connection one form and curvature two form.

Finally, in Section 5 we apply the above introduced theory to the physical example of a gas of non interacting particles in a rigid rotating container and we compute the connection coefficients.

This paper represent a first attempt to study these generalized exponential families and their differential geometric properties. Many important questions are left non answered and need to be treated in future studies. In particular it would be important to study the global structure of the parameter sets $\mathcal{Z}$ and U so that to consider non trivial fibrations of the parameter space. Moreover, since the theory allows so, it would be interesting to see if there are examples which are not related to statistical thermodynamic systems where the fibration π is still meaningful.

2. Exponential families with external parameters

Let X be the (discrete or continuous) state space of the system, $U \subset \mathbb{R}^b$ be the external parameter space and

$$h_\alpha : X \times U \to \mathbb{R}, \quad h = h_\alpha(x, u), \quad \alpha = 1, \ldots, a \tag{4}$$

be the observable functions. The parameter $\theta = (\theta_1, \ldots, \theta_a)$ below are called canonical or natural parameters. Borrowing from the language of thermodynamics, we call the parameters u external parameters. An exponential family is the following probability density depending on (θ, u)

$$p_\theta(x) = e^{-\theta \cdot h(x,u) - \psi(\theta,u)}, \quad \int_X p_\theta dx = 1 \tag{5}$$

where $\theta \cdot h$ denotes scalar product in $\mathbb{R}^a$ and we have defined the free entropy ψ and the partition function Z as

$$\psi(\theta, u) = \ln Z(\theta, u) = \ln \int_X e^{-\theta \cdot h(x,u)} dx \tag{6}$$

We ask that Z be finite and this defines a set of feasible values $z = (\theta, u)$

$$\mathcal{Z} = \{(\theta, u) \in \mathbb{R}^{a+b} \; : \; \int_X e^{-\theta \cdot h(x,u)} dx < +\infty\} \tag{7}$$

which we assume to be open. It is known that for every fixed u the set $\mathcal{Z}$ is open and convex in θ (see [10],[11]). Note that $Z(0, u) = \text{meas}(X)$ hence $\theta = 0$ is a feasible value if and only if X has finite measure. In the following we denote the component of vector θ as θ_α where $\alpha, \beta = 1, \ldots, a$; the component of u as u_k where $k, m = 1, \ldots b$ and the component of $z = (\theta, u)$ as z_i where $i, j = 1, \ldots, a + b$.

2.1. Statistical models and Fisher matrix

We start by recalling the definition of a regular statistical model, see [4], [6]. Let $p_z(x)$ be a probability density defined on state space X depending on finitely many parameters $z \in \mathcal{Z} \subset \mathcal{R}^d$. We introduce the set

$$\mathcal{S} = \{p_z = p(x, z) \quad : \quad z \in \mathcal{Z}\} \subset L^1(X) \tag{8}$$

and we ask that

1) (injectivity) the map $f : \mathcal{Z} \to \mathcal{S}, z \mapsto f(z) = p_z$ is one to one and

2) (regularity) the d functions defined on X

$$p_i(x, z) = \frac{\partial p}{\partial z_i}(x, z), \qquad i = 1, \ldots, d$$

are linearly independent as functions on X for every $z \in \mathcal{Z}$.

This assure that $\mathcal{S}$ is a regular statistical model. The inverse $\varphi : \mathcal{S} \to \mathcal{Z}, \; \varphi(p_z) = z$ of the map f, which exists by 1), defines a global coordinate system for $\mathcal{S}$. It is convenient to introduce the so called log-likelihood

$$l = \ln p \tag{9}$$

and the score basis $l_i = \partial l / \partial z_i$. Note that since $l_i = (1/p)p_i$ the function p_i and l_i are proportional, therefore regularity condition 2) above holds if and only if the elements of the score basis are independent as functions over X. The element of the Fisher matrix are defined as follows

$$g_{ij}(z) = \langle l_i l_j \rangle = \int_X \frac{\partial l}{\partial z_i} \frac{\partial l}{\partial z_j} p dx. \tag{10}$$

If we assume, as we will always do, that we can exchange the operations of integration and differentiation, we have the equivalent definition

$$g_{ij}(z) = \langle l_i l_j \rangle = -\langle \frac{\partial l_i}{\partial z_j} \rangle. \tag{11}$$

The Fisher matrix is symmetric and positive definite therefore it defines a Riemannian metric on $\mathcal{Z}$. (see [6], p.24). In fact we have

$$g_{ij}v_iv_j = \langle l_il_jv_iv_j\rangle = \langle (l_iv_i)^2\rangle = 0 \quad \Leftrightarrow \quad l_iv_i = 0 \quad \Leftrightarrow \quad v_i = 0 \quad \forall\, i \tag{12}$$

since the score vectors l_i are linearly independent over X. Note also (see [6]) that g is invariant with respect to change of coordinates in the state space X and covariant (as a order 2 tensor) with respect to change of coordinates in the parameter space.

In the case of an exponential family (5) depending on natural and external parameters the score basis vectors are, see below (20) and (24) for the explicit computation,

$$l_\alpha = \frac{\partial \ln p}{\partial \theta_\alpha} = -h_\alpha - \frac{\partial \psi}{\partial \theta_\alpha} = \langle h_\alpha\rangle - h_\alpha \tag{13}$$

and

$$l_k = \frac{\partial \ln p}{\partial u_k} = -\theta_\alpha \frac{\partial h_\alpha}{\partial u_k} - \frac{\partial \psi}{\partial u_k} = \theta_\alpha(\langle \frac{\partial h_\alpha}{\partial u_k}\rangle - \frac{\partial h_\alpha}{\partial u_k}) = \theta_\alpha L_{\alpha k}. \tag{14}$$

Note that $\langle l_\alpha\rangle = 0$ and $\langle l_k\rangle = 0$ because $\langle L_{\alpha k}\rangle = 0$. Moreover, one can always assume that $\langle h_\alpha\rangle = 0$ and $\langle \partial h_\alpha/\partial u_k\rangle = 0$ therefore the regularity condition 2) above holds if and only if the $a+b$ functions

$$h_\alpha(x,u), \qquad \theta_\alpha \frac{\partial h_\alpha}{\partial u_k}(x,u) \tag{15}$$

are linearly independent over X. We introduce the more compact notation $\partial_i f = \partial f/\partial z_i$. The elements of the Fisher matrix (10) can be detailed as follows: using (13)

$$g_{\alpha\beta} = \langle l_\alpha l_\beta\rangle = \Big\langle (\langle h_\alpha\rangle - h_\alpha)(\langle h_\beta\rangle - h_\beta)\Big\rangle = cov(h_\alpha, h_\beta) \tag{16}$$

also we have from (14)

$$g_{\alpha k} = \langle l_\alpha l_k\rangle = \Big\langle (\langle h_\alpha\rangle - h_\alpha)\theta_\beta(\langle \frac{\partial h_\beta}{\partial u_k}\rangle - \frac{\partial h_\beta}{\partial u_k})\Big\rangle = \theta_\beta cov(h_\alpha, \partial_k h_\beta) \tag{17}$$

and

$$g_{km} = \langle l_k l_m\rangle = \Big\langle \theta_\alpha(\langle \frac{\partial h_\alpha}{\partial u_k}\rangle - \frac{\partial h_\alpha}{\partial u_k})\theta_\beta(\langle \frac{\partial h_\beta}{\partial u_m}\rangle - \frac{\partial h_\beta}{\partial u_m})\Big\rangle = \theta_\alpha\theta_\beta cov(\partial_k h_\alpha, \partial_m h_\beta). \tag{18}$$

It is useful to introduce a block representation of the symmetric $(a+b)$-dimensional Fisher matrix g as

$$g(z) = \begin{pmatrix} \mathcal{A} & \mathcal{M} \\ \mathcal{M}^T & \mathcal{B} \end{pmatrix} \tag{19}$$

where

$$\mathcal{A}_{\alpha\beta} = g_{\alpha\beta}, \quad \mathcal{M}_{\alpha k} = g_{\alpha k}, \quad \mathcal{B}_{km} = g_{km}.$$

2.2. Expression of Fisher matrix elements using the free entropy function

It is well known that n-order partial derivatives of ψ with respect to the canonical parameters give the n-order moments of the random variables h. By direct computation one has

$$\partial_\alpha \psi = \frac{\partial \psi}{\partial \theta_\alpha} = \frac{\partial}{\partial \theta_\alpha} \ln \int_X e^{-\theta\cdot h} dx = \frac{1}{Z}\int_X e^{-\theta\cdot h}(-h_\alpha)dx = -\int_X p h_\alpha dx = -\langle h_\alpha\rangle \tag{20}$$

and for the second order

$$\partial_\beta \partial_\alpha \psi = -\frac{\partial}{\partial \theta_\beta} \langle h_\alpha \rangle = -\int_X h_\alpha \frac{\partial p}{\partial \theta_\beta} dx = \int_X h_\alpha p (h_\beta + \frac{\partial \psi}{\partial \theta_\beta}) dx = \langle h_\alpha h_\beta \rangle - \langle h_\alpha \rangle \langle h_\beta \rangle \tag{21}$$

where

$$\langle h_\alpha \rangle \langle h_\beta \rangle - \langle h_\alpha \rangle \langle h_\beta \rangle = cov(h_\alpha, h_\beta) \tag{22}$$

is the covariance between the observables h_α and h_β. Hence by comparing (16) and (21) we find the following well known relation (see e.g. [14], [16]) between element of the Fisher matrix g and partial derivatives of the free entropy ψ with respect to canonical parameters θ

$$\mathcal{A}_{\alpha\beta} = g_{\alpha\beta} = \partial_\beta \partial_\alpha \psi. \tag{23}$$

We now compute the the relation between the elements of the Fisher matrix and the partial derivative of ψ with respect to external parameters u. The results are

$$\partial_k \psi = \frac{\partial \psi}{\partial u_k} = \frac{\partial}{\partial u_k} \ln \int_X e^{-\theta \cdot h} dx = \frac{1}{Z} \int_X e^{-\theta \cdot h} (-\theta_\beta \frac{\partial h_\beta}{\partial u_k}) dx = -\theta_\beta \langle \partial_k h_\beta \rangle \tag{24}$$

The right-hand side term is interpreted as an entropic force term in statistical mechanics (see [5] and [8],[13]). For the the second order derivative of the free entropy we have

$$\begin{aligned} \partial_k \partial_\alpha \psi &= -\frac{\partial}{\partial u_k} \langle h_\alpha \rangle = -\int_X (h_\alpha \frac{\partial p}{\partial u_k} + p \frac{\partial h_\alpha}{\partial u_k}) dx = \int_X [h_\alpha p (\theta_\beta \frac{\partial h_\beta}{\partial u_k} + \frac{\partial \psi}{\partial u_k}) + p \frac{\partial h_\alpha}{\partial u_k}] dx \\ &= \int_X [h_\alpha p (\theta_\beta \frac{\partial h_\beta}{\partial u_k} - \langle \theta_\beta \frac{\partial h_\beta}{\partial u_k} \rangle) + p \frac{\partial h_\alpha}{\partial u_k}] dx = \theta_\beta \Big(\langle h_\alpha \partial_k h_\beta \rangle - \langle h_\alpha \rangle \langle \partial_k h_\beta \rangle \Big) - \langle \partial_k h_\alpha \rangle \end{aligned}$$

therefore

$$\partial_k \partial_\alpha \psi = \theta_\beta cov(h_\alpha, \partial_k h_\beta) - \langle \partial_k h_\alpha \rangle \tag{25}$$

and hence by comparing (17) and (25) we get

$$\mathcal{M}_{\alpha k} = g_{\alpha k} = \partial_k \partial_\alpha \psi + \langle \partial_k h_\alpha \rangle. \tag{26}$$

Furthermore

$$\begin{aligned} \partial_k \partial_m \psi &= -\frac{\partial}{\partial u_m} \langle \theta_\alpha \frac{\partial h_\alpha}{\partial u_k} \rangle = -\theta_\alpha \frac{\partial}{\partial u_m} \int_X p \frac{\partial h_\alpha}{\partial u_k} dx = -\theta_\alpha \int_X (\frac{\partial p}{\partial u_m} \frac{\partial h_\alpha}{\partial u_k} + p \frac{\partial^2 h_\alpha}{\partial u_k \partial u_m}) dx \\ &= \theta_\alpha \int_X [\frac{\partial h_\alpha}{\partial u_k} p (\theta_\beta \frac{\partial h_\beta}{\partial u_m} + \frac{\partial \psi}{\partial u_m}) - p \frac{\partial^2 h_\alpha}{\partial u_k \partial u_m}] dx = \theta_\alpha \int_X [\frac{\partial h_\alpha}{\partial u_k} p (\theta_\beta \frac{\partial h_\beta}{\partial u_m} - \langle \theta_\beta \frac{\partial h_\beta}{\partial u_k} \rangle) - p \frac{\partial^2 h_\alpha}{\partial u_k \partial u_m}] dx \\ &= \theta_\alpha \theta_\beta \Big(\langle \partial_k h_\alpha \partial_m h_\beta \rangle - \langle \partial_k h_\alpha \rangle \langle \partial_m h_\beta \rangle \Big) - \theta_\alpha \langle \partial_k \partial_m h_\alpha \rangle \end{aligned}$$

that is

$$\partial_k \partial_m \psi = \theta_\alpha \theta_\beta cov(\partial_k h_\alpha, \partial_m h_\beta) - \theta_\alpha \langle \partial_k \partial_m h_\alpha \rangle \tag{27}$$

hence by comparing (18) and (27) we get

$$\mathcal{B}_{km} = g_{km} = \partial_k \partial_m \psi + \theta_\alpha \langle \partial_k \partial_m h_\alpha \rangle. \tag{28}$$

We see that, unlike the case of natural parameters θ the element of Fisher matrix corresponding to mixed or external parameters do not coincide with second order derivatives of the free entropy.

2.3. *Entropy and energy variation formulae*

Let us compute the Shannon entropy for an exponential family. This can be interpreted ([5]) as the physical entropy of the statistical system defined on X when the the information on the system is given by the average values of the observables for fixed values of the parameters u. We have thus

$$S(\theta,u) = -\int_X p_\theta \ln p_\theta dx = -\int_X p_\theta(-\theta\cdot h - \psi)dx = \psi + \theta_\alpha\langle h_\alpha\rangle = \psi - \theta_\alpha \frac{\partial\psi}{\partial\theta_\alpha} \tag{29}$$

We now compute the differential (one-form) of S as

$$dS = \frac{\partial S}{\partial\theta_\beta}d\theta_\beta + \frac{\partial S}{\partial u_k}du_k. \tag{30}$$

Therefore from (29) and (23) we have

$$\frac{\partial S}{\partial\theta_\beta} = (\frac{\partial\psi}{\partial\theta_\beta} - \frac{\partial\psi}{\partial\theta_\beta}) - \theta_\alpha\frac{\partial^2\psi}{\partial\theta_\alpha\partial\theta_\beta} = -\theta_\alpha cov(h_\alpha, h_\beta) = -\theta_\alpha g_{\alpha\beta} \tag{31}$$

and from (29),(24) and (25)

$$\begin{aligned} \frac{\partial S}{\partial u_k} &= \frac{\partial\psi}{\partial u_k} - \theta_\alpha\frac{\partial^2\psi}{\partial\theta_\alpha\partial u_k} = -\langle\theta_\beta\frac{\partial h_\beta}{\partial u_k}\rangle - \theta_\alpha\Big(\theta_\beta cov(h_\alpha, \frac{\partial h_\beta}{\partial u_k}) - \langle\frac{\partial h_\alpha}{\partial u_k}\rangle\Big) \\ &= -\theta_\alpha\theta_\beta cov(h_\alpha, \partial_k h_\beta) = -\theta_\alpha g_{\alpha k} \end{aligned}$$

hence the infinitesimal entropy variation can be written as, see (19)

$$-dS = \theta_\alpha(g_{\alpha\beta}d\theta_\beta + g_{\alpha k}du_k) = \theta_\alpha(\mathcal{A}_{\alpha\beta}d\theta_\beta + \mathcal{M}_{\alpha k}du_k). \tag{32}$$

For comparison, we write here the average energy variation formula. From (20) and (26) one has

$$\begin{aligned} -d\langle h_\alpha\rangle &= d(\partial_\alpha\psi) = \partial_\alpha\partial_\beta\psi d\theta_\beta + \partial_\alpha\partial_k\psi du_k \\ &= g_{\alpha\beta}d\theta_\beta + (g_{\alpha k} - \langle\partial_k h_\alpha\rangle)du_k \end{aligned}$$

so that using again (19)

$$-d\langle h_\alpha\rangle = \mathcal{A}_{\alpha\beta}d\theta_\beta + \mathcal{M}_{\alpha k}du_k - \langle\partial_k h_\alpha\rangle du_k. \tag{33}$$

We will give a geometric interpretation of these formulae in Section 4 using the theory developed in the next Section 3.

3. Review of Ehresmann connections

For this Section we have followed essentially [15]. On a smooth fibration $\pi : M \longrightarrow N$, where M, N are smooth manifolds, with $\dim M = m$, $\dim N = n$, the set $VM = \ker T\pi$ of the vectors that project onto the null space of TN is an integrable subbundle of TM.

An *Ehresmann connection* on $\pi : M \longrightarrow N$ is the assignment of a distribution HM transversal to VM, so that $HM \oplus VM = TM$. The elements of HM are the horizontal vectors; since $T\pi$ restricted to HM is an isomorphism, it has a fiberwise defined inverse, the horizontal lift: $\text{hor} : T_{\pi(z)}N \longrightarrow T_zM$, $\text{hor}(X) \in H_zM$. Let $X = X^h + X^v$ be the splitting of a vector in T_zM into its horizontal and vertical component. The projection on VM with respect to the horizontal subspace defines the vector–valued connection one-form

$$\omega : TM \longrightarrow VM, \qquad \omega(z)(X) = X^v, \tag{34}$$

whose kernel is the horizontal distribution. The assignment of an horizontal distribution, of an horizontal lift operator or of a connection one-form are equivalent ways to define a connection on $\pi : M \longrightarrow N$.

The *curvature* of the connection is the VM–valued two-form obtained by restricting the exterior derivative of ω to the horizontal distribution:

$$\Omega(X,Y) = d\omega(X^h, Y^h). \tag{35}$$

If we extend the vectors X, Y to vector fields $X, Y \in \Gamma(M)$, and we use Cartan's formula

$$d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]),$$

we get the equivalent expression for the curvature $\Omega(X,Y) = -\omega([X^h, Y^h])$ that shows that the curvature exactly measures the failure of the horizontal distribution to be integrable. Moreover, the curvature relates the Lie brackets of vector fields X, Y on the base manifold N with the Lie bracket of their horizontal lifts, by the formula

$$\Omega(\mathrm{hor}X, \mathrm{hor}Y) = [\mathrm{hor}X, \mathrm{hor}Y]_M - \mathrm{hor}[X,Y]_N. \tag{36}$$

Again, we find that if the curvature is vanishing the horizontal distribution, spanned by vectors of the type hor$X$, is *involutive* hence integrable. Next we give the local expressions of connection and curvature in a fibered chart. Notice that since every foliation is locally a fibration, the following relations hold locally for a foliation. Let $z = (x,y)$ be a fibered chart on $U \subset M$, $\pi(x,y) = y$. Then the vertical space is, $\alpha = 1, \ldots, a = \dim M - \dim N$,

$$V_z U = \ker\ T_z\pi = \mathrm{span}\{\frac{\partial}{\partial x^\alpha}\} \tag{37}$$

and the connection one-form ω is:

$$\omega = \omega^\alpha \otimes \frac{\partial}{\partial x^\alpha}, \quad \omega^\alpha = dx^\alpha + A^\alpha_l(z) dy^l. \tag{38}$$

The $A^\alpha_l(z)$ are the connection's coefficients. The horizontal vectors have the coordinate expression

$$X \in HM \Leftrightarrow \omega(X) = 0, \Leftrightarrow X^l = X^l(\frac{\partial}{\partial y^l} - A^\alpha_l \frac{\partial}{\partial x^\alpha})$$

while the horizontal lift of a base vector $Y = Y^l \frac{\partial}{\partial y^l} \in T_{\pi(z)}N$ has the form

$$(\mathrm{hor}\, Y)^l(z) = Y^l(y)(\frac{\partial}{\partial y^l} - A^\alpha_l \frac{\partial}{\partial x^\alpha}). \tag{39}$$

The connection's curvature can be expressed as: $X, Y \in \Gamma(U)$

$$\Omega(X,Y) = \Omega^\alpha(X,Y) \otimes \frac{\partial}{\partial x^\alpha} = (\Omega^\alpha_{lk} dy^l \otimes dy^k) \otimes \frac{\partial}{\partial x^\alpha} \tag{40}$$

where

$$\Omega^\alpha_{kl} = A^\alpha_{l,k} - A^\alpha_{k,l} - A^\alpha_{l,\beta} A^\beta_k + A^\alpha_{k,\beta} A^\beta_l. \tag{41}$$

Since N is a submanifold, the coordinate base vector fields are Lie-commuting, that is $[\frac{\partial}{\partial y_l}, \frac{\partial}{\partial y_m}] = 0$, formula (36) gives the important result

$$\Omega_{lk} = \Omega(\mathrm{hor}\frac{\partial}{\partial y_l}, \mathrm{hor}\frac{\partial}{\partial y_k}) = [\mathrm{hor}\frac{\partial}{\partial y_l}, \mathrm{hor}\frac{\partial}{\partial y_k}]. \tag{42}$$

See Appendix A for the case of principal bundles.

3.1. Parallel transport equation

Let $\gamma : [0, T] \to N$ be a smooth path in the base manifold and let $z_0 \in \pi^{-1}(\gamma(0))$. The parallel transport equation is the following ODE for the horizontal lift vector field

$$\dot{z} = \frac{dz}{dt} = \mathrm{hor}(\dot{\gamma}), \quad z(0) = z_0 \tag{43}$$

with local expression

$$\dot{x}_\alpha = -A^l_\alpha(x, \gamma)\dot{\gamma}_l, \qquad \dot{y}_l = \dot{\gamma}_l \tag{44}$$

The connection is called *complete* if the parallel transport equation has a solution defined on the whole $[0, T]$. See Appendix B for some results assuring the completeness in the above introduced framework and expecially in case the horizontal distribution is integrable.

3.2. Frames associated to an Ehresmann connection

Let $\varphi \in GL(m, R)$, $\varphi_{ir} = \varphi_{ir}(z)$ be an invertible matrix and define a new base of $T_z M$ as

$$\mathbf{u}_r = \varphi_{ir}\partial_i, \qquad i, r = 1, \ldots, m.$$

Using elementary linear algebra, the relation between the components of a vector $v = w_r \mathbf{u}_r = v_i \partial_i$ with respect to the two frames is as follows,

$$v_i = \varphi_{ir} w_r, \qquad w_r = (\varphi^{-1})_{ri} v_i. \tag{45}$$

If $\mathrm{d}z_j$ is the dual base of the coordinate base, $\mathrm{d}z_j(\partial_i) = \delta_{ij}$, then the dual base of the $\{\mathbf{u}_r\}_{r=1,\ldots,m}$ is given by the n one-forms

$$\zeta_r(z) = (\varphi^{-1})_{ri}\mathrm{d}z_i, \qquad \zeta_r(\mathbf{u}_s) = \delta_{rs}. \tag{46}$$

The frame $\{u_r\}_{r=1,\ldots,m}$ is called *nonholonomic* frame if the Lie bracket of at least two base vectors is non vanishing. Indeed, by Frobenius theorem, one can show that if $[\mathbf{u}_r, \mathbf{u}_s] = 0$ for all r, s then there exists a change of chart $z' = \eta(z)$ such that $u_r = \partial/\partial z'_r$ and $\varphi_{ri} = \partial\eta_r/\partial z_i$. If (M, g) is a Riemannian manifold, the metric g has the following form in a non coordinate frame

$$g_{rs} = g(\mathbf{u}_r, \mathbf{u}_s) = g(\varphi_{ir}\partial_i, \varphi_{js}\partial_j) = \varphi_{ir}\varphi_{js}g_{ij} \tag{47}$$

and the metric g_{rs} can be used for raising or lowering indices in the non coordinate frame.

Our first application of the above outlined theory is to define a frame associated to an Eheresmann connection on a given fiber bundle. Give an arbitrary vector $v \in T_z M$, $v = v_i \partial_i$, we can consider the splitting of the same vector with respect to $T_z M = V_z M \oplus H_z M$ into its vertical and horizontal part

$$v = \omega(v) + \mathrm{hor}(T\pi(z)v). \tag{48}$$

The frame, see (38) and (39), associated to the Ehresmann connection is defined as

$$\mathbf{u}_\alpha = \frac{\partial}{\partial x^\alpha}, \quad \alpha = 1, \ldots, a \quad \mathbf{u}_l = \mathrm{hor}(\frac{\partial}{\partial y^l}), \quad l = 1, \ldots n \tag{49}$$

It is a nonholonomic frame if the horizontal subbundle is a non integrable distribution. In particular, componentwise, see (38) and (39), if $v = X^\alpha \partial_\alpha + Y^l \partial_l$ we get

$$v = X^\alpha \partial_\alpha + Y^l \partial_l = w^\alpha \mathbf{u}_\alpha + w^l \mathbf{u}_l$$

where

$$w^\alpha = \omega^\alpha(v), \quad w^l = Y^l = (T\pi(v))^l.$$

It is a straightforward computation to show that, from (45) the matrices $\varphi \in GL(m)$ have the following block form

$$\varphi = \begin{pmatrix} \mathbb{I}_a & -A \\ 0 & \mathbb{I}_n \end{pmatrix}, \qquad \varphi^{-1} = \begin{pmatrix} \mathbb{I}_a & A \\ 0 & \mathbb{I}_n \end{pmatrix}. \tag{50}$$

Moreover, from (46), the dual base of the associated frame is

$$\zeta^\alpha = \omega^\alpha,\ \alpha = 1, \ldots a, \qquad \zeta^l = \mathrm{d}y^l,\ l = 1, \ldots, n. \tag{51}$$

Using again a block representation for the matrices, the Riemannian metric g has the following expression in the associated frame

$$g = \begin{pmatrix} \mathcal{A} & \mathcal{M} \\ \mathcal{M}^T & \mathcal{B} \end{pmatrix}, \qquad G = \varphi^T g \varphi = \begin{pmatrix} \mathcal{A} & \mathcal{M} - \mathcal{A}A \\ \mathcal{M}^T - A^T\mathcal{A} & \mathcal{C} \end{pmatrix} \tag{52}$$

where

$$\mathcal{C} = \mathcal{B} + A^T\mathcal{A}A - A^T\mathcal{M} - \mathcal{M}^T A.$$

We now specialize the above relations to the important case where the horizontal distribution M_z is defined to be the g-orthogonal of V_zM. Referring tho the above introduced block representation of the metric g we ask that every $X^h \in H_zM,\ X^h = (-A^\alpha_l Y^l, Y^l)$ be orthogonal to all $X^v \in V_zM$, $X^v = (w, 0)$. As a consequence

$$g(X^v, X^h) = w \cdot (-\mathcal{A}AY + \mathcal{M}Y) = 0 \quad \forall\, w \quad \Leftrightarrow \quad A = \mathcal{A}^{-1}\mathcal{M}. \tag{53}$$

In the orthogonal splitting case the metric G has the simpler form

$$G = \varphi^T g \varphi = \begin{pmatrix} \mathcal{A} & 0 \\ 0 & \mathcal{K} \end{pmatrix}, \quad \text{where} \quad \mathcal{K} = \mathcal{B} - \mathcal{M}^T\mathcal{A}\mathcal{M} \tag{54}$$

or from (48)

$$g(z)dz \otimes dz = \mathcal{A}(z)\omega(\cdot) \otimes \omega(\cdot) + \mathcal{K}(z)dy \otimes dy \tag{55}$$

The connection one-form (38) becomes from (53)

$$\omega^\alpha = dx^\alpha + A^\alpha_l(z)dy^l, \qquad \text{where} \quad A^\alpha_l = (\mathcal{A}^{-1}\mathcal{M})^\alpha_l \tag{56}$$

and it is called mechanical connection in the control theory for mechanical systems, where g is the kinetic energy of a mechanical system. In the special case of a deformable body where the fiber of the fiber bundle is $SO(3)$, then the analog of $g_{\alpha\beta}$ is called the locked inertia tensor.

4. Fibered exponential statistical models

Let us consider again the fibration (3) between natural and external parameters

$$\pi : \mathcal{Z} \longrightarrow U, \qquad \pi(\theta, u) = u$$

The theory exposed in Section 3 for a general fibration $\pi : M \longrightarrow N$, $\pi(x, y) = y$ is applied here for $x = \theta$ and $y = u$; the natural parameters θ are coordinates on the fiber $\pi^{-1}(u)$ and the external parameters u are coordinates in the base U. It is too complex and far from our objectives to investigate the global geometric and topologic nature of the set $\mathcal{Z} \subset \mathbb{R}^{a+b}$ introduced in (7) of Section 1. Therefore we suppose that $\mathcal{Z}$ and U are open set of an euclidean space and that they give the description of the fibration in a local chart.

The Fisher matrix g defines a Riemannian metric in $\mathcal{Z}$. Let us define the horizontal subspace to be the g-orthogonal subspace of the vertical subspace. Using the block representation (19) -see also (52)- for the Fisher metric g, the connection coefficients in case of a g-orthogonal connection are

$$A_k^\alpha = (\mathcal{A}^{-1}\mathcal{M})_l^\alpha = (g^{-1})_{\alpha\beta} g_{\beta k}.$$

Now the formula (32) that gives the variation of entropy dS can be rewritten as follows

$$-dS = \theta_\alpha (g_{\alpha\beta} d\theta_\beta + g_{\alpha k} du_k) = \theta_\alpha \mathcal{A}_{\alpha\beta}(d\theta_\beta + A_k^\alpha du_k) = \theta_\alpha \mathcal{A}_{\alpha\beta}\omega^\beta$$

therefore the one form dS can be rewritten as

$$dS = -\theta_\alpha \mathcal{A}_{\alpha\beta}\omega^\beta = -\langle \theta, \omega \rangle_{\mathcal{A}}. \tag{57}$$

where $\langle\ ,\ \rangle_{\mathcal{A}}$ denotes the scalar product between vertical vectors in the orthogonal splitting of the metric g in (54).

A first consequence of the above introduced framework is the following: let $z(t) = (\theta(t), \gamma(t))$ for $t \in [0,1]$ be the horizontal lift of a path in the external parameters $u = \gamma = \gamma(t)$. Then the variation of the entropy $S(\theta, u)$ along the horizontal path z is from (57)

$$\Delta S = \int_0^1 \frac{dS}{dt} dt = \int_0^1 dS(z)\dot{z} dt = \int_0^1 \langle \theta, \omega(\dot{z}) \rangle_{\mathcal{A}} dt = 0 \tag{58}$$

since $\dot{z}$ is an horizontal vector. Then the horizontal lift of a path in the mechanical parameters represent an isentropic ($\Delta S = 0$) evolution for the system.

A second consequence is that the average energy variation formula (33) can be rewritten as

$$-d\langle h_\alpha \rangle = \mathcal{A}_{\alpha\beta} d\theta_\beta + \mathcal{M}_{\alpha k} du_k - \langle \partial_k h_\alpha \rangle du_k = \mathcal{A}_{\alpha\beta}\omega^\beta - \langle \partial_k h_\alpha \rangle du_k.$$

If we define $\langle h_\alpha \rangle$ as the α-type average energy, $P_k^\alpha = -\langle \partial_k h_\alpha \rangle$ as the (α, k)-type pressure and

$$-\,dQ_\alpha = \mathcal{A}_{\alpha\beta}\omega^\beta, \qquad (\mathcal{A}_{\alpha\beta} = g_{\alpha\beta}) \tag{59}$$

the α-type infinitesimal heat exchange, then we have the following equality between one-forms

$$d\langle h_\alpha \rangle = dQ_\alpha - P_\alpha^k du_k \tag{60}$$

which is a generalized form of the first principle of thermodynamics. Using (57) we obtain

$$dS = -\theta_\alpha \mathcal{A}_{\alpha\beta}\omega^\beta = \theta_\alpha dQ_\alpha = \theta_\alpha \left(d\langle h_\alpha \rangle + P_\alpha^k du_k \right) \tag{61}$$

which is the first principle of thermodynamics in the form of Gibbs. The above geometrical interpretation of the dS and dQ one forms is contained also in [5] but without using the language of fibrations, which we consider particularly effective for the following developments.

Remark. In Souriau Lie Groups thermodynamics, (see [10], [9] for a recent in-depth analysis) the observables h_α (first integrals of the system) take value in the dual $\mathcal{G}^*$ of the Lie algebra of a Lie group G acting on the system phase space while the canonical parameters θ called generalized temperatures are element of $\mathcal{G}$ and $\theta \cdot h$ is substituted by $\langle h, \theta \rangle$ where $\langle\ ,\ \rangle$ denotes the duality between $\mathcal{G}$ and $\mathcal{G}^*$. In our setting, in case of a principal bundle with typical fiber G, VM can be identified with $\mathcal{G}$ and $g_{\alpha\beta}\omega^\beta \in \mathcal{G}^*$ (see Appendix A). Note however the difference in the notion of heat between Souriau Lie Group thermodynamics and here: since Souriau theory does not consider external parameters u, then the first principle of Thermodynamics (60) reads $d\langle h_\alpha \rangle = dQ_\alpha$. Accordingly, in Souriau theory the generalized or geometric α-type heat is defined as $Q_\alpha = \langle h_\alpha \rangle$.

Remark. Let us consider the induced fibration

$$\mathcal{S} \xrightarrow{\varphi} \mathcal{Z} \xrightarrow{\pi} U, \qquad \tilde{\pi} = \pi \circ \varphi$$

where φ is the coordinate system for $\mathcal{S}$. Fibers of $\tilde{\pi}$ are the probability densities p_z where $z = (\theta, u)$ and u is fixed. So they represent standard exponential families. Let

$$\sigma : T\mathcal{Z} \longrightarrow T\mathcal{S}, \qquad \sigma(z, \partial_i) = (f(z), s(\partial_i) = (p_z, l_i)$$

be the fibered map that sends the canonical base $\partial_i = \partial/\partial z_i$ into the score base $l_i = \partial_i \ln p$. Then, by definition of scalar product in $T\mathcal{S}$

$$\langle l_i, l_j \rangle = \int_X p_z l_i l_j dx = g_{ij}(z)$$

the orthogonal frame of $T_z\mathcal{Z}$ (see (49))

$$\mathbf{u}_\alpha = \partial_\alpha, \quad \alpha = 1, \ldots, a \quad \mathbf{u}_k = \text{hor}(\partial_k), \quad k = 1, \ldots, b$$

is sent by σ into the orthogonal frame of $T_{p_z}\mathcal{S}$

$$l_\alpha = s(\partial_\alpha), \qquad \lambda_k = s(\text{hor}(\partial_k)) = s(\partial_k - A_k^\beta \partial_k) = l_k - A_k^\beta l_\beta$$

Indeed, since $A = \mathcal{A}^{-1}\mathcal{M}$,

$$\langle l_\alpha, \lambda_k \rangle = \langle l_\alpha, l_k \rangle - A_k^\beta \langle l_\alpha, l_\beta \rangle = \mathcal{M}_{\alpha k} - (\mathcal{A}^{-1})_{\beta\gamma}\mathcal{M}_{\gamma k}\mathcal{A}_{\alpha\beta} = \mathcal{M}_{\alpha k} - \mathcal{M}_{\alpha k} = 0.$$

Then we have the orthogonal splitting $T\mathcal{S} = V\mathcal{S} \oplus H\mathcal{S}$ of the tangent space where

$$V\mathcal{S} = \text{span}\{l_\alpha\}, \qquad H\mathcal{S} = \text{span}\{\lambda_k\}.$$

4.1. Parallel transport for the g-orthogonal connection

Let us consider the parallel transport equation (43) above in the case of exponential models. Let γ be a closed path i.e. $\gamma(0) = \gamma(1)$, in the base manifold representing an externally imposed cyclical evolution in the external parameters. Then we have the parallel transport equation (44)

$$\dot{\theta}_\alpha = -A_k^\alpha \dot{\gamma}_k = -(\mathcal{A}^{-1})_{\alpha\beta}\mathcal{M}_{\beta k}\dot{\gamma}_k.$$

Equivalently, the integration along the horizontal lift $\tilde{z}$ of γ of the vector-valued temperature one form $d\theta = -Adu$ gives

$$\Delta\theta(\gamma) = \theta(1) - \theta(0) = -\int_\gamma A = -\int_0^1 (\mathcal{A}^{-1}\mathcal{M})(\tilde{z}(t))dt$$

The last integral represents the integration of the one form A over the path u and it is invariant with respect to reparameterizations of the path. The quantity $\Delta\theta(\gamma)$ is called the *holonomy* of the path. Note that if the holonomy of the path is nonzero we may have for an horizontal lift of a closed path γ in base manifold see (58)

$$\Delta S(\gamma) = \int_\gamma \theta_\alpha dQ_\alpha = 0, \qquad \text{but} \quad \Delta\theta(\gamma) \neq 0.$$

5. Gas in a rotating container

This system is studied in [7] and [9]. See also [11] for the treatment of this system in Souriau' Lie Group thermodynamic framework. We consider a gas of N non interacting point particles contained in a rigid container $B \subset \mathbb{R}^3$ with a fixed point O. Let $\mathcal{E}'$ be the inertial frame and $\mathcal{E}$ the frame attached to B, both having the origin in O, and let ω be the angular velocity of the motion of $\mathcal{E}$ with respect to $\mathcal{E}'$. We denote with a prime the quantities expressed in the inertial frame. If we neglect the interaction of the particles with the container walls, the Lagrangian of the system computed in the inertial and moving frame are respectively

$$L' = \frac{1}{2}\sum_i m_i v_i'^2 = \frac{1}{2}\sum_i m_i (v_i + \omega \times r_i)^2 = L.$$

We want to compute the Hamiltonian of the system in the rotating frame. We have, with a little abuse of notation

$$p_i = \frac{\partial L}{\partial v_i} = m_i(v_i + \omega \times r_i), \quad \tilde{v}_i = \frac{p_i}{m_i} - \omega \times r_i$$

and

$$H(r,p) = \sum_i p_i \cdot \tilde{v}_i - L(\tilde{v}, r) = \frac{1}{2}\sum_i \frac{p_i^2}{m_i} - \omega \cdot \sum_i r_i \times p_i = \frac{1}{2}\sum_i \frac{p_i^2}{m_i} - \omega \cdot J(r,p) \tag{62}$$

where $J(r,p) = \sum_i r_i \times p_i$ is the angular momentum of the gas with respect to the origin O in the rotating frame. Here the angular velocity is an external parameter which can be varied at will. If we suppose that the average energy of the system is known, then the probability distribution that describes the microscopic state of the system (r,p) in $X^N = B^N \times \mathbb{R}^{3N}$ is the exponential density

$$\rho(r,p) = \frac{e^{-\theta H(r,p,\omega)}}{Z(\theta,\omega)}, \qquad Z(\theta,\omega) = \int_{X^N} drdp\, e^{-\theta H(r,p,\omega)}. \tag{63}$$

Since the particle are non interacting, it is straightforward to derive that the probability density ρ is the product of N single particle probability densities which differs only by the mass m_i

$$\rho(r,p) = \prod_{i=1}^N \rho_i(r_i, p_i)$$

Note that the log-likelihood of the N-particle system can be written as

$$l(r,p) = \ln \rho = \sum_{i=1}^N \ln \rho_i(r_i,p_i) = \sum_{i=1}^N l^{(i)}(r_i,p_i)$$

and one can show (see [6]) that since the N random variables (r_k, p_k) are independent

$$g_{ij} = \langle l_i l_j \rangle_\rho = N \langle l_i^{(k)} l_j^{(k)} \rangle_{\rho_k} = N g_{ij}^{(k)}.$$

From now on we will consider the probability density of a single particle which can be derived from (63) for $N = 1$ and $m_1 = m$. The single particle partition function and free entropy are

$$Z(\theta,\omega) = \int_X drdp\, e^{-\theta(\frac{p^2}{2m} - \omega \cdot r \times p)}, \qquad \psi = \ln Z \tag{64}$$

Note that this system fits in the scheme of Section 2 for $a = 1, b = 3, X = B \times \mathbb{R}^3, x = (r, p)$ and $u = \omega$, $U = \mathbb{R}^3$. For the sake of simplicity we will denote the external parameters with ω instead of u. The only observable is the energy

$$h(r,p,\omega) = h_0(r,p) - \omega \cdot J(r,p) = \frac{p^2}{2m} - \omega \cdot r \times p \tag{65}$$

where $h_0(r,p)$ is the kinetic energy term. Note that the function h is linear in the external parameters ω and this make some terms in the formulae (27) and (28) of Section 2 to vanish. We have $\partial_k h = J_k$ and

$$l = \ln p = -\theta h(r,p,\omega) - \psi(\theta,\omega).$$

The score vectors are

$$l_\theta = \frac{\partial l}{\partial \theta} = -h - \partial_\theta \psi = \langle h \rangle - h,$$

$$l_k = \frac{\partial l}{\partial \omega_k} = -\theta \partial_k h - \partial_k \psi = -\theta(\langle J_k \rangle - J_k), \quad k = 1,2,3$$

which are linearly independent since $p^2/2m$ and $J_k, k = 1,2,3$ are linearly independent. The elements of the Fisher matrix are computed from (23), (26) and (28)

$$\mathcal{A} = g_{\theta\theta} = \partial^2_\theta \psi = cov(h,h) = var(h)$$

$$\mathcal{M}_k = g_{\theta k} = \theta cov(h, J_k) = \partial_k \partial_\theta \psi + \langle J_k \rangle$$

$$\mathcal{B}_{km} = g_{km} = \theta^2 cov(J_k, J_m) = \partial_k \partial_m \psi$$

so the expression of the connection one form coefficient (56) is

$$A_k = \mathcal{A}^{-1} \mathcal{M}_k = g^{-1}_{\theta\theta} g_{\theta k} = \frac{\theta cov(h, J_k)}{var(h)} = \frac{\partial_k \partial_\theta \psi + \langle J_k \rangle}{var(h)}$$

5.1. Alternate form of the Fisher metric

We can provide an alternate expression for the elements of Fisher matrix g as follows. By adding $\pm m(\omega \times r)^2$ one can rewrite the single particle energy (65) as

$$h = \frac{p^2}{2m} - \omega \cdot r \times p = \frac{m}{2}[(\frac{p}{m} - \omega \times r)^2 - (\omega \times r)^2]$$

hence using Fubini theorem to iterate the integral the single particle partition function (64) becomes

$$\begin{aligned} Z(\theta,\omega) &= \int_X drdp\, e^{-\theta h(r,p,\omega)} = \int_X drdp\, e^{-\theta \frac{m}{2}[(\frac{p}{m} - \omega\times r)^2 - (\omega \times r)^2]} \\ &= \int_B dr\, e^{\theta \frac{m}{2}(\omega \times r)^2} \int_{\mathbb{R}^3} dp\, e^{-\theta \frac{m}{2}(\frac{p}{m} - \omega \times r)^2} \\ &= \int_B dr\, e^{\theta \frac{m}{2}(\omega\times r)^2} \int_{\mathbb{R}^3} dp\, e^{-\theta \frac{p^2}{2m}} \\ &= \int_B dr\, e^{-\theta U(\omega,r)} \left(\frac{2\pi m}{\theta}\right)^{\frac{3}{2}} = Z_B(\theta,\omega)\left(\frac{2\pi m}{\theta}\right)^{\frac{3}{2}} \end{aligned} \tag{66}$$

where in the last line we have introduced the centrifugal potential

$$U(\omega, r) = -\frac{m}{2}(\omega \times r)^2 = -\frac{1}{2}\omega \cdot I(r)\omega$$

and $I(r) = m(r^2\mathbb{I} - r \otimes r)$ is the inertia tensor of the particle with respect to the origin O. From the above expression (66) for $Z(\theta, \omega)$ we see that the set of feasible values is

$$\mathcal{Z} = \{(\theta, \omega) \ : \ Z(\theta, \omega) < +\infty\} = (0, +\infty) \times \mathbb{R}^3. \tag{67}$$

Note also that $Z(\theta, \omega)$ coincides with the partition function of a particle described by the Hamiltonian

$$h(r, p) = \frac{p^2}{2m} + U(\omega, r) = \frac{p^2}{2m} - \frac{m}{2}(\omega \times r)^2. \tag{68}$$

We can compute the free entropy from (66) as

$$\psi = \ln Z = \ln Z_B(\theta, \omega)\left(\frac{2\pi m}{\theta}\right)^{\frac{3}{2}} = -\frac{3}{2}\ln\theta + \frac{3}{2}\ln 2\pi m + \tilde{\psi} \tag{69}$$

where we have introduced the free entropy

$$\tilde{\psi}(\theta, \omega) = \ln Z_B(\theta, \omega) = \ln \int_B dr\, e^{-\theta U(\omega, r)}. \tag{70}$$

Note that $\tilde{\psi}$ is the free entropy of the probability density on the configuration space $B \subset \mathbb{R}^3$

$$\tilde{p}(\theta, \omega; r) = e^{-\theta U(\omega, r) - \tilde{\psi}(\theta, \omega)}, \qquad \int_B \tilde{p} dr = 1.$$

We denote with $\langle\ \rangle_B$ the expected values taken with respect to the probability $\tilde{p}$. It is not difficult to show using Fubini theorem as in (66) that for a function only of the particle position $f(r)$ one has

$$\langle f \rangle = \int_X drdp\, e^{-\theta h - \psi} f(r) = \int_B dr e^{-\theta U(\omega, r) - \tilde{\psi}(\theta, \omega)} f(r) = \langle f \rangle_B. \tag{71}$$

5.2. Computation of Fisher matrix and connection one form

Starting form the free entropy (69) we have

$$\partial_\theta \psi = -\frac{3}{2\theta} + \partial_\theta \tilde{\psi} = -\frac{3}{2\theta} - \langle U \rangle_B = -\frac{3}{2\theta} - \frac{1}{2}\omega \cdot \langle I \rangle_B \omega$$

and introducing the gradient $\nabla = (\partial_1, \partial_2, \partial_3)$

$$\nabla_\omega \psi = \nabla_\omega \tilde{\psi} = -\theta \langle \nabla_\omega U \rangle_B = \theta \langle I \rangle_B \omega = \theta \langle J \rangle_B$$

If we make the standard identification $\theta = 1/k_B T$ where T is the temperature and k_B is the Boltzmann constant we can recognize that the above expression has the form of an entropic force. Moreover with similar computations we have

$$\partial_k \partial_\theta \psi = \partial_k \partial_\theta \tilde{\psi} = -\langle \partial_k U \rangle_B - \theta cov_B(U, \partial_k U).$$

Now we can compute the Fisher matrix elements. From (23) we have

$$\mathcal{A} = g_{\theta\theta} = \partial_\theta^2 \psi = \frac{3}{2\theta^2} + \partial_\theta^2 \tilde{\psi} = \frac{3}{2\theta^2} - var_B(U)$$

and from (26)

$$\mathcal{M}_k = g_{\theta k} = \partial_k \partial_\theta \psi + \langle J_k \rangle = -\langle \partial_k U \rangle_B - \theta cov_B(U, \partial_k U) + \langle J_k \rangle_B$$

Moreover (28) becomes

$$\mathcal{B}_{km} = g_{km} = \partial_m \partial_k \psi = \partial_m \partial_k \tilde{\psi} = -\theta \langle \partial_m \partial_k U \rangle_B + \theta^2 cov_B(\partial_k U, \partial_m U)$$

The connection one form is now

$$A_k(\theta, \omega) = \mathcal{A}^{-1} \mathcal{M}_k = g_{\theta\theta}^{-1} g_{\theta k} = \frac{-\langle \partial_k U \rangle_B - \theta cov_B(U, \partial_k U) + \langle J_k \rangle_B}{\frac{3}{2\theta^2} - var_B(U)}$$

and the curvature coefficients can be computed in principle from (41).

5.3. Curved exponential family

We can obtain a explicit expression for the free entropy $\tilde{\psi}$ in (70) if we specify its shape and restrict the possible motions of the rigid container. We suppose that B is a rigid cylinder of length L and radius R and that the attached frame $\mathcal{E}$ has its z axis parallel to the longitudinal axis of B and origin O fixed in the geometric center of B. Moreover we suppose that $\mathcal{E}$ rotates with angular velocity $\omega = \omega(\lambda) = \lambda \hat{z}$ where $\hat{z}$ is the unit vector of the z axis in the attached *and* inertial frame. In this case (keeping the same symbol for U as a function of λ and ω)

$$U(\lambda, r) = U(\omega(\lambda), r) = -\frac{m}{2}(\omega \times r)^2 = -\frac{m\lambda^2}{2}(\hat{z} \times r)^2$$

Using cylindrical coordinates (ρ, ϕ, z) the configurational partition function Z_B can be explicitly computed as

$$Z_B(\theta, \lambda) = \int_B e^{-\theta U(\lambda, r)} dr = \int_0^{2\pi} d\phi \int_{-L/2}^{L/2} dz \int_0^R d\rho \rho e^{\frac{\theta m \rho^2 \lambda^2}{2}} = \frac{2\pi L}{\theta m \lambda^2}(e^{\frac{mR^2}{2}\theta\lambda^2} - 1)$$

and the free entropy is the explicit function of the parameters (θ, λ)

$$\begin{aligned} \psi(\theta, \lambda) &= -\frac{3}{2}\ln\theta + \ln Z_B = -\frac{3}{2}\ln\theta + \ln\frac{1}{\theta\lambda^2} + \ln\frac{2\pi L}{m}(e^{\frac{mR^2}{2}\theta\lambda^2} - 1) \\ &= -\frac{5}{2}\ln\theta - 2\ln\lambda + \ln\frac{2\pi L}{m}(e^{\frac{mR^2}{2}\theta\lambda^2} - 1) \end{aligned} \tag{72}$$

The energy of the system see (68) depends on the external parameter λ

$$h(\lambda, \rho, p) = h(\omega(\lambda), \rho, p) = \frac{p^2}{2m} + U(\lambda, r) = \frac{p^2}{2m} - \frac{m}{2}\rho^2\lambda^2 \tag{73}$$

and the log-likelihood is now depending in the two parameters (θ, λ)

$$l(\theta, \lambda, \rho, p) = -\theta h(\rho, p, \lambda) - \psi(\theta, \lambda)$$

so now the statistical model is spanned by the two score vectors l_θ, l_λ and this define a submanifold of the initial statistical manifold called curved exponential model in the statistical literature (see [4]). Now the element of the Fisher matrix can be computed explicitly but the fibration $\pi : \mathbb{R} \times [0, +\infty) :\longrightarrow [0, +\infty)$, $\pi(\theta, \lambda) = \lambda$ has a one dimensional simply connected base manifold so that the Ehresmann connection is always trivially flat. Using (73) we have immediately

$$\partial_\lambda h = -m\rho^2\lambda, \qquad \partial_\lambda^2 h = -m\rho^2.$$

We have from (72) (here Csch is the hyperbolic cosecant)

$$\mathcal{A} = g_{\theta\theta} = \partial^2_\theta \psi = \frac{5}{2\theta^2} - \frac{m^2 R^4 \lambda^4}{16} \mathrm{Csch}(\frac{mR^2}{4}\theta\lambda^2)^2$$

and from (26) and the result (71)

$$\mathcal{M} = g_{\theta\lambda} = \partial_\lambda \partial_\theta \psi + \langle \partial_\lambda h \rangle = \partial_\lambda \partial_\theta \psi - m\lambda \langle \rho^2 \rangle_B$$

where

$$\partial_\lambda \partial_\theta \psi = \frac{mR^2 \lambda e^{\frac{mR^2}{2}\theta\lambda^2}(2e^{\frac{mR^2}{2}\theta\lambda^2} - mR^2\lambda^2\theta - 2)}{2(e^{\frac{mR^2}{2}\theta\lambda^2} - 1)}$$

while from (28) and (71) we have

$$\mathcal{B} = g_{\lambda\lambda} = \partial^2_\lambda \psi + \theta \langle \partial^2_\lambda h \rangle = \partial^2_\lambda \psi - m \langle \rho^2 \rangle_B.$$

where

$$\partial^2_\lambda \psi = \frac{2 + e^{mR^2\theta\lambda^2}(2 + \lambda^2 mR^2\theta) - e^{\frac{mR^2}{2}\theta\lambda^2}(4 + \lambda^2 mR^2\theta(1 + \lambda^2 mR^2\theta))}{\lambda^2(e^{\frac{mR^2}{2}\theta\lambda^2} - 1)}$$

The above formulae are completely determined when we compute (here Erfi is the error function)

$$\langle \rho^2 \rangle_B = R + \frac{2R\lambda - \frac{\sqrt{2\pi}\mathrm{Erfi}(\frac{R\lambda\sqrt{m\theta}}{\sqrt{2}})}{\sqrt{m\theta}}}{2\lambda(e^{\frac{mR^2}{2}\theta\lambda^2} - 1)}.$$

We have given the as far as possible explicit expression of the Fisher metric and connection coefficient in this example to show that the theory exposed in the previous Sections can be applied to physical systems. However the explicit expression are rather heavy and not easy to read.

6. Conclusions

In this paper we have investigated a generalized form of exponential families, whose introduction is motivated by equilibrium statistical mechanics. The first aim in the paper is the computation of the associated Fisher metric g, a notion which is relevant for the study of the differential geometric properties (Riemann metric, Christoffel symbols) and the study of phase transitions (Riemannian curvature). A second step is the introduction of a trivial fibration of the parameter space over the external parameters space and of the g-orthogonal sub bundle which define an Ehresmann connection on the fibration. This geometric framework is well suited for studying the evolution of the thermodynamic system when the evolution in the external parameters is given. It turns out that the horizontal lift of a path in the external parameters space represent an isentropic evolution of the whole system. This is a first simple consequence of the presence of the Fisher metric and the fibration which could be generalized once one considers non trivial fibrations. Some of the possible applications of the theory introduced in this paper are contained in the fully worked out example of Section 5 and in the Appendices.

Appendix A. The principal bundle case

We particularize the notion of Ehresmann connection to the important class of *principal* fiber bundles, where the fibration is the one defined by the set of orbits of a smooth group action. Suppose that a Lie group G acts freely and properly on the left on M and that, to every $z \in M$, $z \mapsto Gz$ is an immersion, so that $\pi : M \longrightarrow M/G$ is a principal bundle. On a principal bundle, the group action induces an Ad–equivariant isomorphism $\sigma : \mathcal{G} \longrightarrow \Gamma(M, VM)$, $\quad \sigma(\xi) := \xi_M(z)$, where ξ_M is the infinitesimal vector field associated to $\xi \in \mathcal{G}$. A *principal* connection on the bundle is the assignment of a Ehresmann connection compatible with the group action, i.e. satisfying

$$T\Phi_g(H_z M) = H_{gz}M \quad \forall z \in M, \quad \forall g \in G. \tag{A1}$$

Moreover, the related connection one form $\tilde{\omega}$ is G–invariant and it defines an equivariant $\mathcal{G}$–valued connection one form by setting $\omega = \sigma^{-1} \circ \tilde{\omega}$. In a principal bundle, the curvature, defined as above, is Ad–equivariant and, unlike the Ehresmann case, it satisfies the *structure equation*

$$\Omega(X,Y) = \mathrm{d}\omega(X,Y) - [\omega(X), \omega(Y)]. \tag{A2}$$

Now consider a local trivialization $z = (y, g)$ of the bundle, where g are the coordinates on the fiber isomorphic to G. By equivariance of ω, the local expression of the connection become

$$\omega(y,g)(\dot{y},\dot{g}) = \mathrm{Ad}_g(\xi + A(y)\dot{y}) \tag{A3}$$

where $\xi = g^{-1}\dot{g} \in T_e G = \mathcal{G}$ is the left translation to the origin of $\dot{g} \in T_g G$ and A_l^α are the connection's coefficients with respect to a chosen basis $\{e_\alpha\}$ of $\mathcal{G}$. It is important to notice that now the connection's coefficients are represented by functions constant on the fibers.

Appendix B. The flat connection case

The following definition introduces a sufficient condition for a Ehresmann connection to be complete.

Definition A1. *let $\pi : M \to N$ be a smooth fibration with (M, g) a Riemannian manifold. The metric g is called bundle-like with respect to the fibration π if in the adapted local coordinates (x, y) for the above splitting (55) of g one has $\mathcal{K} = \mathcal{K}(y)$. As a consequence, the horizontal lift hor : $T_{\pi(z)}N \longrightarrow H_z M$ is an isometry.*

It is not difficult to prove that if the metric is bundle-like, then the connection is complete. Indeed, if by absurdum the interval where the maximal solution $z_m(\cdot)$ is defined is $[0, \tau)$ with $\tau < T$, since the length of $z_m((0, \tau))$ coincides with the length of $\gamma((0, \tau))$ which is finite, then the maximal solution is contained in a compact set of M and by a classical result of ODE theory, the solution of z_m can be extended to a whole neighborhood of τ.

The connection is *flat* when the horizontal distribution is integrable. In this case the horizontal lift of a path lies on a single leaf of the horizontal foliation; moreover, the parallel transport depends only on the *homotopy* class of the path, therefore the parallel transport operator H_γ defines a group homomorphism of the fundamental group $\pi_1(y_0, N)$ into $\mathrm{Diff}(\pi^{-1}(y_0))$, the group of diffeomorphisms of the fiber.

Appendix C. Computation of Christoffel symbols

Since the Fisher metrics is a Riemannian metric, one can compute its Christoffel symbols of first kind

$$\Gamma_{ij,w} = \frac{1}{2}\left(\partial_i g_{jw} + \partial_j g_{iw} - \partial_w g_{ij}\right) \tag{A4}$$

and the derivatives of the Fisher metric can be computed e.g. using the formula

$$-\partial_w g_{ij} = \langle \partial_i \partial_j l_w \rangle + \langle (\partial_i l_j) l_w \rangle \tag{A5}$$

so that the Christoffel symbols have the equivalent expression

$$2\Gamma_{ij,w} = \langle (\partial_i l_j) l_w \rangle - \langle (\partial_w l_j) l_i \rangle - \langle (\partial_w l_i) l_j \rangle - \langle \partial_i \partial_j l_w \rangle \tag{A6}$$

In the case of a generalized exponential family, from (13) and (14) we derive immediately that

$$\partial_\beta l_\alpha = -\partial_\alpha \partial_\beta \psi$$

and

$$\partial_k l_\alpha = \partial_k h_\alpha - \partial_k \partial_\alpha \psi \tag{A7}$$

and

$$\partial_m l_k = -\theta_\alpha \partial_m \partial_k h_\alpha - \partial_m \partial_k \psi = -\partial_m \partial_k (\theta \cdot h) - \partial_m \partial_k \psi \tag{A8}$$

Then plugging the above expression in (A6) we get the following table of coefficients

$$\begin{aligned}
2\Gamma_{\alpha\beta,\gamma} &= -\langle \partial_\alpha \partial_\beta l_\gamma \rangle \\
2\Gamma_{\alpha\beta,k} &= \langle (\partial_k h_\beta) l_\alpha + (\partial_k h_\alpha) l_\beta - \partial_\alpha \partial_\beta l_k \rangle = 2\Gamma_{\beta\alpha,\gamma} \\
2\Gamma_{\alpha k,\beta} &= \langle (\partial_k h_\beta) l_\alpha - \partial_k h_\alpha) l_\beta - \partial_\alpha \partial_k l_\beta \rangle = 2\Gamma_{k\alpha,\beta} \\
2\Gamma_{\alpha k,m} &= \langle (\partial_m h_\alpha) l_k - (\partial_k h_\alpha) l_m + (\partial_m \partial_k (\theta \cdot h)) l_\alpha - \partial_\alpha \partial_k l_m \rangle = 2\Gamma_{k\alpha,m} \\
2\Gamma_{km,\alpha} &= \langle (\partial_k h_\alpha) l_m + (\partial_m h_\alpha) l_k - (\partial_m \partial_k (\theta \cdot h)) l_\alpha - \partial_m \partial_k l_\alpha \rangle = 2\Gamma_{mk,\alpha} \\
2\Gamma_{km,n} &= \langle (\partial_n \partial_m (\theta \cdot h)) l_k + (\partial_n \partial_k (\theta \cdot h)) l_m - (\partial_k \partial_m (\theta \cdot h)) l_n - \partial_k \partial_m l_n \rangle = 2\Gamma_{mk,n}
\end{aligned}$$

References

1. Fisher, R. A. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London*. Series A, Containing Papers of a Mathematical or Physical Character **1922**
2. Rao, C. R. Information and the accuracy attainable in the estimation of statistical parameters. In *Breakthroughs in statistics* Springer, New York, NY, 1992; pp. 235-247.
3. Amari, S. Nagaoka, H. *Methods of Information Geometry* AMS and Oxford University Press, 2000
4. Amari, S. *Information geometry and its applications* Vol. 194. Springer, 2016.
5. Jaynes, E.T. Brandeis Lectures. In *Papers on Probability, Statistics and Statistical Physics*; Rosenkrantz, E.D.; Springer, Dordrecht; pp. 39-76
6. Calin, O., Udriste, C. *Geometric modeling in probability and statistics* Springer, Berlin, 2014
7. Balian, R. *From Microphysics to Macrophysics. Methods and Applications of Statistical Physics*. Volume I Theoretical and Mathematical Physics, 2006
8. Crooks, G. E. Fisher information and statistical mechanics. *Tech. Rep.* **2011**.
9. Marle, C-M. From tools in symplectic and Poisson geometry to J-M. Souriau theories of statistical mechanics and thermodynamics *Entropy* **2016**18.10 : 370.
10. Barbaresco, F. Geometric theory of heat from Souriau Lie groups thermodynamics and Koszul Hessian geometry: Applications in information geometry for exponential families. *Entropy* **2016** 18.11 386.
11. Souriau, J-M. *Structure of dynamical systems: a symplectic view of physics* Vol. 149. Springer Science & Business Media, 2012.
12. Roos, N. Entropic forces in Brownian motion. *American Journal of Physics* **2014** 82.12 : 1161-1166.
13. Rubinstein, M., Ralph H. C. *Polymer physics* Vol. 23. New York: Oxford University Press, 2003.
14. Prokopenko, M., et al. Relating Fisher information to order parameters. Physical Review E **2011** 84.4 041116.
15. Marsden, J.E., Montgomery, R., Ratiu T.S. *Reduction, symmetry, and phases in mechanics* Vol. 436. American Mathematical Soc., 1990.
16. Janke, W., Johnston, D. A., Kenna,R. Information geometry and phase transitions. *Physica A: Statistical Mechanics and its Applications* **2004** 181-186.